CLAIMS

1.   A system for data mining from one or more data sources comprising: a source of data comprising one or more domains of information;

an Object-Relationship Database comprising objects from the one or more domains of information; and

a knowledge discovery engine where relationships between two or more integrated objects are identified, retrieved, grouped, ranked, filtered and numerically evaluated.

2.   The system of claim 1, wherein the source is one or more databases containing textual information.

3.   The system of claim 1, wherein the source is one or more databases containing numerical information.

4.   The system of claim 1, wherein the relationships between the two or more objects are identified as direct or indirect.

5.   The system of claim 4, wherein the relationships between the two or more integrated objects are ranked based on the relative strength of the relationship between direct and indirect objects.

6.   The system of claim 1, wherein the relationships are set into categories selected from the group consisting of positive, negative, physical and logical associations.

7.   The system of claim 1, wherein the domains of information comprise parcels of data as information as text, symbol, numerals and combinations thereof.

8.   The system of claim 1, wherein the system is at least partially automated.

9.   The system of claim 1, wherein the knowledge discovery engine filters the two or more integrated objects by lexical processing.

10.   The system of claim 1, wherein the Object-Relationship Database (ORD) is created using a method comprising the steps of:

compiling one or more data source objects;

adding the synonyms of the data source objects; and

grouping the information in the one or more data source into an object-relationship database.

11.   The system of claim 10, further comprising a database of lexical variants from a data source.

12.   The system of claim 11, wherein the system further comprises a program for scanning the object-relationship database with the database of lexical variants to add synonyms;

13.   The system of claim 12, wherein the system comprises a program for checking the object-relationship database for errors.

14.   The system of claim 10, wherein the ORD creation method further comprises the step of increasing processing efficiency by assigning each database object a unique numeric ID and storing adirectional relationships by lowest ID first.

15.   The system of claim 1, wherein an object is retrieved from unstructured text, structured data, a list, a table, a phrase, a paragraph, an abstract, a program, a manual, a text book, a reference book, treatise, a lab notebook, a letter, a memo, an email, a table of contents, index, a magazine, an article, scientific literature, a patent, a patent application, an international application, a webpage, a spreadsheet, a URL, or relational database, and combinations thereof.

16.   The system of claim 15, wherein the object is selected from the group consisting of from the group consisting of gene, protein, chemical compound, small molecule, drugs, diseases, clinical phenotype, and other identifiers selected from the group consisting of ChemID, MeSH, FDA, locuslink, GDB, HGNC, MeSH, Medline, Snowmed, and OMIM.

17.   The system of claim 10, wherein the ORD creation method further comprises the step of screening out common words.

18.   The system of claim 10, wherein the ORD further comprises the step of identifying capitalizations and patterns for words by accessing a word database.

19.   The system of claim 11, wherein the step of constructing lexical variants further comprises using a synonym database.

20.   The system of claim 10, wherein the step of constructing lexical variants further comprises using an acronym-resolving algorithm.

21.   The system of claim 1, further comprising a graphical user interface for displaying one or more objects.

22.   The system of claim 21, wherein the interface comprise a control element, which can be clicked to display the integrated object derived from the context of the source data.

23.   The system of claim 1, wherein a portion of the Object-Relationship Database is constructed using a method comprising the steps of:
inputting a block of text from the source of data;
extracting information from the source to create a record and
creating one or more arrays to match words in the record against phrases in the object-relationship database.

24.   The system of claim 23, wherein the method further comprises resolving acronyms.

25.   The system of claim 23 or24, wherein the method further comprises parsing the record into sentences and parsing each sentence into words.

26.   The system of claim 23, wherein the information comprises title, abstract, date, and PMID fields.

27.   The system of claim 22, wherein the block of text is selected from the group consisting of a list, a table, a phrase, a paragraph, an abstract, a program, a manual, a text book, a reference book, a lab notebook, a letter, a memo, an email, a table of contents, a magazine, an article, scientific literature, a patent, a patent application, an international application, a webpage, a spreadsheet, a URL, or relational database, and combinations thereof.

28.   The system of claim 27, wherein the block of text is selected from the Physician's Desk Reference.

29.   The system of claim 23, wherein the block of text is given a higher value if the source of the information is considered to have a higher impact than other like sources according to selected

criteria for impact.

30.   A system for relating objects comprising:

an object-relationship database generated from a data source comprising one or more domains of information; and

a knowledge discovery engine that recognizes relationships between objects in a data source, wherein the knowledge discovery engine identifies a one or more cooccurrences of objects within the data source, and identifies implicit relationships between the objects.

31.   The system of claim 30, wherein the knowledge discovery engine generates a comprehensive network of relationships.

32.   The system of claim 31, wherein the knowledge discovery network generates a partial network of relationships.

33.   The system of claim 30, wherein the relationships idenfied are stored in a system database and the system further includes a query module that allows a user to access information about the implicit relationships.

34.   The system of claim 30, wherein the knowledge discovery engine evaluates relationships using one or more statistically bounded network models.

35.   A system for identifying a new indication for a drug comprising:

an object-relationship database generated from a data source comprising one or more domains of information including information relating to the drug; and

a knowledge discovery engine that recognizes meaningful relationships in a data source for the drug, wherein the knowledge discovery engine identifies one or more co-occurrences of objects within the data source and the drug, and

generates a comprehensive network of relationships between objects in the object-relationship database and the drug, wherein at least one relationship identifies a new indication for the drug.

36.   The system of claim 35, wherein the knowledge discovery engine evaluates

relationships using one or more statistically bounded network models.

37.   The system of claim 35, wherein the system further stores shared and implicit relationships in a results database.

38.   A system for identifying a contraindication and/or side-effect for a drug comprising:

an object-relationship database generated from a data source comprising one or more domains of information including information relating to the drug; and

a knowledge discovery engine that recognizes meaningful relationships in the object relationship database, wherein the knowledge discovery engine identifies one or more cooccurrences of objects and a drug in a data source, identifies shared and implicit relationships between objects and the drug, and identifies the likelihood that one or more of the relationships indicates one or more contraindications and/or side-effect of the drug.

39.   The system of claim 38, wherein the knowledge discovery engine generates a comprehensive network of relationships between data in the data source and the drugs, and stores the shared and implicit relationships evaluated by one or more statistically bounded network models.

40.   A system for identifying interactions between at least two drugs comprising:

an object-relationship database generated from a data source comprising one or more domains of information including information relating to the at least two drugs; and

a knowledge discovery engine that recognizes meaningful relationships in the object relationship database, wherein the knowledge discovery engine identifies one or more cooccurrences of objects and drugs in the data source, identifies shared and implicit relationships between objects and the drugs, and identifies the likelihood that co-occurrence of the one or more objects with the at least two drugs indicates an interaction between the at least two drugs. Could also be two genes or a drug and a gene, ie other relationships of value.

41.   The system of claim 40, wherein the knowledge discovery engine generates a comprehensive network of relationships between data in the data source and the drugs and stores the shared and implicit relationships evaluated by one or more statistically bounded network models.

42.   A system for identifying relationships between a chemical compound or biomolecule and a disease comprising:

an object-relationship database generated from a data source comprising one or more domains of information including information relating to the disease and a chemical compound or biomolecule; and

a knowledge discovery engine that recognizes meaningful relationships in the data source for the disease, wherein the knowledge discovery engine: identifies one or more co-occurrences of objects, the disease and/or the chemical compound or biomolecule within the data source, and identifies shared and implicit relationships between the chemical compound or biomolecule and the disease.

43.   The system of claim 42, wherein the knowledge discovery engine generates a comprehensive network of relationships between data in the object-relationship database and the disease and stores the shared and implicit relationships evaluated by one or more statistically bounded network models.

44.   The sysetm of claim 42, wherein the biomolecule is a nucleic acid or protein.

45.   The system of any of claims 1, 30, 35, 38, 40 or 42, further comprising a scanning module comprising a scanner for scanning printed information and generating a data source from the printed information.

46.   The system of any of claims 1, 30, 35, 38, 40 or 42, wherein the system comprises a processor for executing the functions of the knowledge engine.

47.   The system of claim 46, further comprising a computer readable medium for storing the object-relationship database.

48.   The system of claim 47, further comprising a client/server architecture wherein at least two functions of the system are distributed in a server and at least one client computer connectable to the network.

49.   The system of claim 48, wherein the system comprises a program for accessing one or more data sources.

50.   The system of claim 48, wherein the object relationship database is dynamic, and adds new objects from the one or more data sources to the database.

51.   The system of claim 50, wherein the system recomputes an object network when new objects are added from the one or more data sources.

52.   The system of claim 51, wherein the system further comprises an engine for monitoring recomputation results; and wherein the system re-evaluates relationships between objects.

53.   The system of claim 48, wherein the database is downloadable to the at least one client computer.

54.   The system of claim 48, wherein the database (network) is stored in memory of the server computer and the at least one client can access the database by communicating with the server.

55.   The system of any of claims 1, 30, 35, 38, 40 or 42, wherein the system further comprises a results and analysis database, wherein the results and analysis database comprises:information relating to a query regarding an object relationship and results of the query.

56.   The system of claim 55, wherein the results and analysis database further comprises a record of comprising information relating to an interpretation of the results.

57.   The system of claim 55, wherein the results and analysis database further comprises data validating the results.

58.   The system of any of claims 1, 30, 35, 38, 40 or 42, wherein the system further comprises an application program for executing a computer code comprising instructions for

ranking relationships.

59.   The system of claim 58, wherein the computer code includes instructions for a system processor to generate a linear or nonlinear gouping of individual ranking factors.

60.   The system of claim 59, wherein each individual ranking factor is associated with a coefficient that weights each term.

61.   The system of claim 60, wherein weight is determined by one or more of the following factors: the source of the data source; the date on which the data source was published; the ratio of the observed frequency of co-occurrence of objects to the expected frequency of co-occurrence of objects; the name of the author associated with the data source; the name of the institution associated with the data source; and the frequency of co-occurrence of objects in different data sources.

62.   A method for data mining from a data source comprising one or more domains of knowledge comprising the steps of:
        obtaining or accessing a data source;
        generating an Object-Relationship Database comprising objects from the data source data; and
        identifying the strength of direct and implicit relationships in the Object-Relationship database.

63.   The method of claim 62, wherein data in the data source source is searched for co-occurrences of objects in the source of data, and objects are retrieved from the data source for storing in the Object-Relationship database based on the co-occurences.

64.   The method of claim 61, wherein the data is selected from the group consisting of unstructured text, structured data, a list, a table, a phrase, a paragraph, an abstract, a program, a manual, a text book, a reference book, treatise, a lab notebook, a letter, a memo, an email, a table of contents, index, a magazine, an article, scientific literature, a patent, a patent application, an international application, a webpage, a spreadsheet, a URL, or relational database, and combinations thereof.

65. The method of claim 63 wherein relationships are ranked according to their strength.

66. The method of claim 63, wherein strength is determined by one or more of the following factors: the source of the data source; the date on which the data source was published; the ratio of the observed frequency of co-occurrence of objects to the expected frequency of co-occurrence of objects; the name of the author associated with the data source; the name of the institution associated with the data source; and the frequency of co-occurrence of objects in different data sources.

67. A method for relating objects comprising the steps of:
generating an object-relationship database generated from a data source comprising one or more data sources, or accessing the object-relationship database; and
identifying implicit relationships between objects using a knowledge discovery engine; and
determining the strength of the relationships.

68. The method of claim 61, wherein the frequency of co-occurrences of objects within the datasource is determined.

69. The method of claim 61, wherein the knowledge discovery engine generates a comprehensive network of relationships to identify the implicit relationships.

70. The method of claim 67, wherein the strength of the relationships are evaluated using one or more statistical bounded network models.

71. A method for identifying a new indication for a drug comprising:
obtaining or accessing an object-relationship database generated from a data source which includes information relating to the drug; and
processing information in the object-relationship database with a knowledge discovery engine that recognizes meaningful relationships, by identifying one or more co-occurrences of objects from the data source; generating a comprehensive network of relationships between objects in the object-relationship database and the drug to identify implicit relationships between the object and the drug, wherein at least one relationship identifies a new indication for the drug.

72.   The method of claim 71, further comprising storing shared relationships evaluated by one or more statistical bounded network models.

73.   A method for identifying a contraindication or side-effect for a drug comprising:

obtaining or accessing an object-relationship database generated from a data source comprising one or more domains of information including information relating to the drug; and

processing information in the object-relationship database with a knowledge discovery engine that recognizes meaningful relationships in the object relationship database, wherein the knowledge discovery engine identifies one or more cooccurrences of objects and a drug in a data source, identifies shared and implicit relationships between objects and the drug, and identifies the likelihood that one or more of the relationships indicates one or more contraindications and/or side-effects of the drug.

74.   A method for identifying interactions between at least two drugs comprising:

obtaining or accessing an object-relationship database generated from a data source comprising one or more domains of information including information relating to the at least two drugs; and

processing information in the object-relationship database with a knowledge discovery engine that recognizes meaningful relationships in the object relationship database, wherein the knowledge discovery engine identifies one or more cooccurrences of objects and drugs in the data source, identifies shared and implicit relationships between objects and the drugs, and identifies the likelihood that co-occurrence of the one or more objects with the at least two drugs indicates an interaction between the at least two drugs.

75.   A method for identifying relationships between a chemical compound or a biomolecule and a disease comprising:

obtaining an object-relationship database generated from a data source comprising one or more domains of information; and

processing information in the object-relationship database using a knowledge discovery engine wherein the knowledge discovery engine: identifies one or more co-occurrences of objects, the disease and/or the chemical compound or biomolecule within the data source, and identifies shared and implicit relationships between the chemical compound or biomolecule and the disease.

76.   A method for creating an Object-Relationship Database (ORD) comprising the steps of:

compiling one or more objects from one or more data sources

grouping the information in the one or more data sources into an object-relationship database;

constructing a database of lexical variants from one or more data sources;

comparing the database of lexical variants to objects in the Object-Relationship Database;

scanning the object-relationship database with the database of lexical variants to add synonyms

assigning each object a unique numeric ID and storing adirectional relationships by lowest ID first; and

checking the object-relationship database for errors.

77.   The method of claim 76, wherein the data sources used to compile the database objects are selected from the group consisting of chemical compounds, small molecules, diseases, phenotypes, genes, proteins, clinical data, drugs, identifiers from ChemID, identifiers from MeSH, identifiers from FDA, identifiers from locuslink, identifiers from GDB, identifiers from HGNC, identifiers from MeSH, identifiers from OMIM.

78.   The method of claim 76, wherein the data sources to compile the database objects include a list, a table, a phrase, a paragraph, an abstract, a program, a manual, a text book, a reference book, a lab notebook, a letter, a memo, an email, a table of contents, a magazine, an article, scientific literature, a patent, a patent application, an international application, a webpage, a spreadsheet, a URL, or relational database, and combinations thereof.

79.   The method of claim 76, wherein one or more data sources or portions of one or more data sources are scanned to extract new objects.

80.   The method of claim 76, wherein the extracting step comprises selecting objects in the context of data from one or more data sources or portions thereof and determining whether the object is included in the Object-Relationship Database.

81.   The method of claim 80, wherein if the object is not included, it is stored in Object-Relationship Database.

82.   The method of claim 80, wherein information relating to whether objects are included in the Object-Relationship Database is displayed on a graphical user interface.

83.   The method of claim 82, wherein the data scanned and selected is also displayed on the graphical user interface.

84.   The method of claim 76, wherein an object in the object relationship database is text, a number or symbol.

85   The method of claim 76, further comprising the step of filtering the objectrelationship database for ambiguous acronyms using a word database.

86.   The method of claim 76, further comprising the step of identifying lexical variants using a synonym database.

87.   The method of claim 76 or 85, further comprising the step of identifying lexical variants using an acronym-resolving algorithm.

88.   The method of claim 76, further comprising the step of providing the object in the context of the text from the source of data in the database.

89.   The method of claim 76, furhter comprising the step of reducing redundancies in the data source.

90.   The method of claim 89, wherein the method of reducing redundancies comprises the steps of:

inputting a block of text from a source;

extracting information from the source to create a record;

parsing the record into sentences; parsing each sentence into words;

creating one or more arrays to match words against phrases in the object-relationship database;

flagging acronyms; and

storing the acronyms in the database of lexical variants.

91.   A method for identifying novel correlative relationships comprising the steps of: identifying
one or more topical clusters from a data source;

compiling a database of objects from one or more topical clusters;

refining the database of objects to reduce redundancies;

scanning the topical set from the data source for co-occurring objects;

identifying co-occurring objects as relationships;

analyzing the identified relationships for statistical relevance with respect to one or
more objects;

creating one or more relationship databases; and

storing the relationships and relationship databases.

92.   The method of claim 91, wherein the step of compiling the database of objects
further comprises the steps of:

creating fields of interest that are grouped together;

identifying databases that house similar groups of information;

preprocessing the database entries into pre-defined formats;

resolving the entries; and

checking for errors to remove uninteresting entries based on a pre-selected criteria.

93.   The method of claim 91, wherein the step of refining the database of objects further
comprises the step of flagging ambiguous acronyms using a word database for lexical variants.

94.   The method of claim 91, wherein the step of refining the database of objects further
comprises the step of scanning a source for the existence of co-occurring objects to reduce
redundancies and create relationships, which comprises the steps of:

inputting a block of text from the source;

extracting data from the block of text;

parsing the data into sentences;

parsing each sentence into words;

putting the words into one or more arrays;

matching the object database for matches against the words from any array; and

determining whether there is a match between the object database and the words from the array.

95.   The method of 94, wherein the step of identifying relationships within the relationship database comprises the steps of:

assigning each object a unique numeric ID; and

storing adirectional relationships by lowest ID first.

96.   The method of claim 94, wherein the step of identifying relationships within the relationship database comprises the steps of:

identifying shared relationships after a user inputs one or more lists of objects for analysis;

compiling all from the one or more lists all the relationships for each object into a single

list;

counting related objects by frequency; and

calculating an expectation value.

97.   The method of claim 85, further comprising the steps of:

excluding shared objects with less than an x% of the total possible connection or less than

a y% of the observed/expected ratio;

identifying implicitly related objects for each shared relationship; and

scoring implicitly related objects by direct observed/expected ratio times the number of unique paths to the implicit object.

98.   The method of claim 97, wherein the user varies the x% of the total possible connection to vary the score of implicit relationships.

99.   The method of claim 97, wherein the user varies the y% of the observed/expected ratio to vary the score of implicit relationships.

100.   The method of claim 97, wherein the correlative relationship is between a drug, chemical compound, small molecule, phenotype, disease, gene, genotype and combinations thereof.

101.   A method of evaluating direct relationships between one or more objects comprising the steps of:

computing an association strength vector between one or more first, second and third objects;

obtaining a source impact score from a database of source impact scores for the one or more objects for the first, second or third objects; and

multiplying the strength vector by the source impact score for one or more of the first, second or third objects.

102.   The method of claim 101, wherein the source impact score is based on the publication from which the one or more objects were obtained.

103.   The method of claim 101, wherein the source impact score is based on the number of times the source of the one or more objects were cited by another source.

103.   The method of claim 101, wherein the source impact score is based on the number of times the source of the one or more objects were cited by a treatise.

104.   The method of claim 101, wherein the source impact score is based on the number of times the source of the one or more objects were cited in one or more textbooks.

105.   The method of claim 101, wherein the source impact score is based on the number of times the source of the one or more objects were cited in a review article.

106.   The method of claim 101, wherein the source impact score is given a score based on its estimated importance and relevance.

107.   The method of claim 101, wherein the source impact score is given a score based on the number of times the source of the one or more objects were published in a peer reviewed journal.

108.   The method of claim 101, wherein a higher impact score implies higher importance and relevance.

109.   A computer program embodied on a computer readable medium for accessing domains of information comprising:

a code segment adapted to contain a source of data comprising one or more domains of

information;

a code segment adapted to maintain an Object-Relationship Database; and

a code segment adapted to contain a knowledge discovery engine where relationships between two or more objects are searched, grouped, ranked, filtered, and retrieved.

110. A computer program embodied on a computer readable medium for creating an Object-Relationship Database (ORD) comprising:

a code segment adapted to compile one or more database objects;

a code segment adapted to group the information in the one or more databases into an object-relationship database;

a code segment adapted to construct a database of lexical variants from one or more databases;

a code segment adapted to scan the object-relationship database with the database of lexical variants to add synonyms; and

a code segment adapted to assign each object a unique numeric ID and storing adirectional relationships by lowest ID first; and

a code segment adapted to check the object-relationship database for errors.

111. A data structure comprising a plurality of candidate compounds for new drug therapy generated by a method comprising the steps of:

accessing a source of data comprising one or more domains of information;

compiling the domains of information into an Object-Relationship Database for integrating objects from the one or more domains of information; and

using a knowledge discovery engine where relationships between two or more integrated objects are identified, retrieved, grouped, ranked, filtered and numerically evaluated.

112. A data structure comprising a plurality of candidate compounds for evaluation generated by a method comprising the steps of:

obtaining an object-relationship database generated from a data source comprising one or more databases of information; and

processing one or more objects using a knowledge discovery engine to recognize meaningful relationships from a data source comprising the steps of:

identifying one or more co-occurrences of objects from the data source; generating a

comprehensive network of relationships; and

storing the shared relationships evaluated by one or more statistical bounded network models, wherein a query is performed on the shared relationships to identify novel relationships from the comprehensive network of relationships.

113.   A system for identifying a previously unidentified use for a compound comprising the steps of:

obtaining an object-relationship database generated from a data source comprising one or more domains of information including information relating to the compound; and processing the information in the data source using a knowledge discovery engine thatrecognizes meaningful relationships between a drug and one or more objects by identifying one or more co-occurrences of objects in a data source; generating a comprehensive network of relationships; and storing the shared relationships evaluated by one or more statistical bounded network models.

114.   A method of treating cardiac hypertrophy comprising the steps of:

identifying a patient in need of therapy for cardiac hypertrophy; and

providing the patient with a pharmaceutically effective amount of a compound identified using the system of claim 1 using a query comprising the term cardiac hypertrophy.

115.   A method of treating cardiac hypertrophy comprising the steps of:

providing a patient in need of the treatment with a therapeutically effective amount of a Chlorpromazine.

116.   A method of treating cardiac hypertrophy comprising the steps of:

providing a patient in need of the treatment with therapeutically effective amount of a Chlorpromazine.

117.   A method of treating cardiac hypertrophy comprising the steps of:

providing a patient in need of the treatment with a therapeutically effective amount of a compound (make another claim for groups of compounds that would be used together in a combination therapy) selected from the group consisting of: compound selected from the group consisting of: Naloxone, Naltrexone,Triiodothyronine, Clonidine, Estrogen, Tamoxifen,

Colchicine, Bradykinin, Omapatrilat, Apstatin, COX-2 selective inhibitor, 5-LOX inhibitor, Thromboxane A2 Receptor Antagonist, Melatonin, Morphine, Warfarin/Heparin, Cortisol,and Methionine.

118.   A method for treating of non-insulin dependent diabetes mellitus (NIDDM) comprising the steps of:

identifying a patient in need of therapy for NIDDM; and

providing the patient with a therapeutically effective amount of a compound identified using the system of claim 1.

119.   A method for treating of non-insulin dependent diabetes mellitus (NIDDM) comprising the steps of:

administering to a patient in need of therapy for NIDDM; a therapeutically effective amount of a compound that increases the methylation of cellular nucleic acids.

120.   A method for treating of non-insulin dependent diabetes mellitus (NIDDM) comprising the steps of:

adminstering to a patient in need of therapy for NIDDM, a therapeutically effective amount of DNA methylation precursors.

121.   A nutritional supplement for an individual at risk for of non-insulin dependent diabetes mellitus (NIDDM) comprising:

one or more DNA methylation precursors at an amount effective to normalize the level of DNA methylation.

122.   A method for treating migraine headaches comprising the steps of:

identifying a patient in need of therapy for a migraine headache; and providing the patient with a therapeutically effective amount of sildenafil .

123.   A method for treating muscular spasms comprising the steps of:
identifying a patient in need of therapy for a muscular spasm; and
providing the patient with a therapeutically effective amount of sildenafil.

124.   A system for automated screening comprising:

a system as described in claim 1, wherein the object relationship database includes objects which are nucleic acid or protein sequences or identifiers of such sequences;

an oligonucleotide selection module that selects nucleic acid sequences based on relationships between objects and genes corresponding to the nucleic acid and/or protein sequences and/or identifiers of the sequences, using the knowledge engine and provides instructions to a DNA-on-chip assembly apparatus to immobilize the selected nucleic acid sequences on a solid support.

125.   The system of claim 124, wherein the instructions are provided to the apparatus via a user of the system.

126.   The system of claim 124, wherein the nucleic acid sequences have been identified by the system as having a correlation to NIDDM.

127.   A method for numerically assigning importance to each relationship identified using the system of claim 1 comprising the steps of:

identifying one or more co-occurrences of objects within one or more topical sets in a domain of information; and

evaluating the probability that one or more co-occurrences of objects represents a meaningful relationship within one or more topical sets.

128.   The method of claim 127, wherein the importance is a function of the number of times two objects are co-mentioned within the topical set in the domain of information.

129   The method of claim 127, wherein the importance is a function of the textual distance between two objects.

130.   The method of claim 127, wherein the importance is based on an external measure of the topical set, wherein the external measure is selected from the group consisting of importance, relevance, and quality.

131.   The method of claim 127, wherein the importance includes an evaluation of one or more co-occurrence patterns over time.

132.   The method of claim 127, wherein a natural language processing engine is used to

identify one or more co-occurrences of objects.

133.   The method of claim 127, wherein contextual information within the topical set is used to assign importance.

134.   The method of claim 133, wherein contextual information within the topical unit of text is used to assign a nature to the relationship.

135.   The method of claim 127, wherein importance is veracity.

136.   A method of finding implicit relationships comprising the steps of

identifying one or more objects directly related to one or more query objects as a set of directly related objects;

identifying one or more objects related to the set of directly related objects as a set of implicitly related objects; and

quantitatively evaluating each implicitly related object to determine a probability that it shares a meaningful relationship with the query object by deriving an importance score and a veracity score.

137.   The method of claim 136, wherein quantitative evaluation further comprises a probability that a statistically similar relationship could be observed by chance.

138.   The method of claim 136, wherein a formula (6) according to

$$P(A \leftrightarrow B_1^n) = \sum_1^n 1 - (1 - \frac{K_A}{Nt}) * (1 - \frac{K_{B1}}{Nt})^n . \qquad (6)$$

is used.

139.   A method of identifying relationships shared by one or more objects in a set comprising a plurality of objects; comprising the steps of:

enumerating a set of objects;

identifying all new objects related to the set from a data source; and

quantitatively evaluating the statistical significance that the new object is related to the set.

140. The method of claim 139, wherein objects that link other objects to the set are identified and used to identify one or more relationships common to the set.

141. The method of claim 139, wherein one or more topical groupings in the set are identified and distinguished from random groupings based on their cohesiveness.

142. The method of claim 139, wherein the new object is added to the set if the statistical significance meets a selected value.

143. The method of claim 139, wherein at least one object corresponds to a biomolecule arrayed on a microarray, a biomolecule that binds to an array, a gene, an expression value of a biomolecule, a phenotype, a disease, a small molecule, a chemical compound, a metabolite, a drug, a therapeutic agent, a candidate gene, an expressed sequence, and combinations thereof.

144. The method of claim 143, wherein the expression value comprises 0 or 1, wherein 0 is not expressed and 1 is expressed.

145. The method of claim 143, the expression value comprises a quantitative measure of expression.

146. The method of claim 143, wherein the set comprises objects which include expression values and the new object comprises an expression value.

147. The method of claim 146, wherein the expression value of the new object is evaluated to determine its relationship to known objects of the set.

148. The method of claim 139, wherein a quantitative evaluation of the probability that the new object shares a meaningful relationship with the set is determined by deriving an importance score and a veracity score.

149. The method of claim 139, wherein quantitative evaluation further comprises a probability that a statistically similar relationship could be observed by chance.

150. A data structure comprising an implicit relationship as set forth in FIGURE 25.

151. A computer program product stored on a computer readable medium comprising program code for executing functions of the system of any of claims 1, 30, 35, 38, 40 or 42, and 124.

152. The method of claim 71, wherein the drug is sildenafil.